BAC sequencing analyses in the hexaploid *Spartina maritima* (Poaceae): Homoeolog divergence and microsynteny in the grass family

Carine Charron[1], Mathieu Rousseau-Gueutin[2], Julie Ferreira de Carvalho[2], Houda Chelaifa[2], Julien Boutte[2], Gaëtan Droc[1], Joëlle Fourment[3], Arnaud Bellec[3], Hélène Berges[3], Julie Poulain[4], Arnaud Couloux[4], Sophie Mangenot[4], Patrick Wincker[4], Armel Salmon[2], Angélique D'Hont[1] and Malika Ainouche[2]

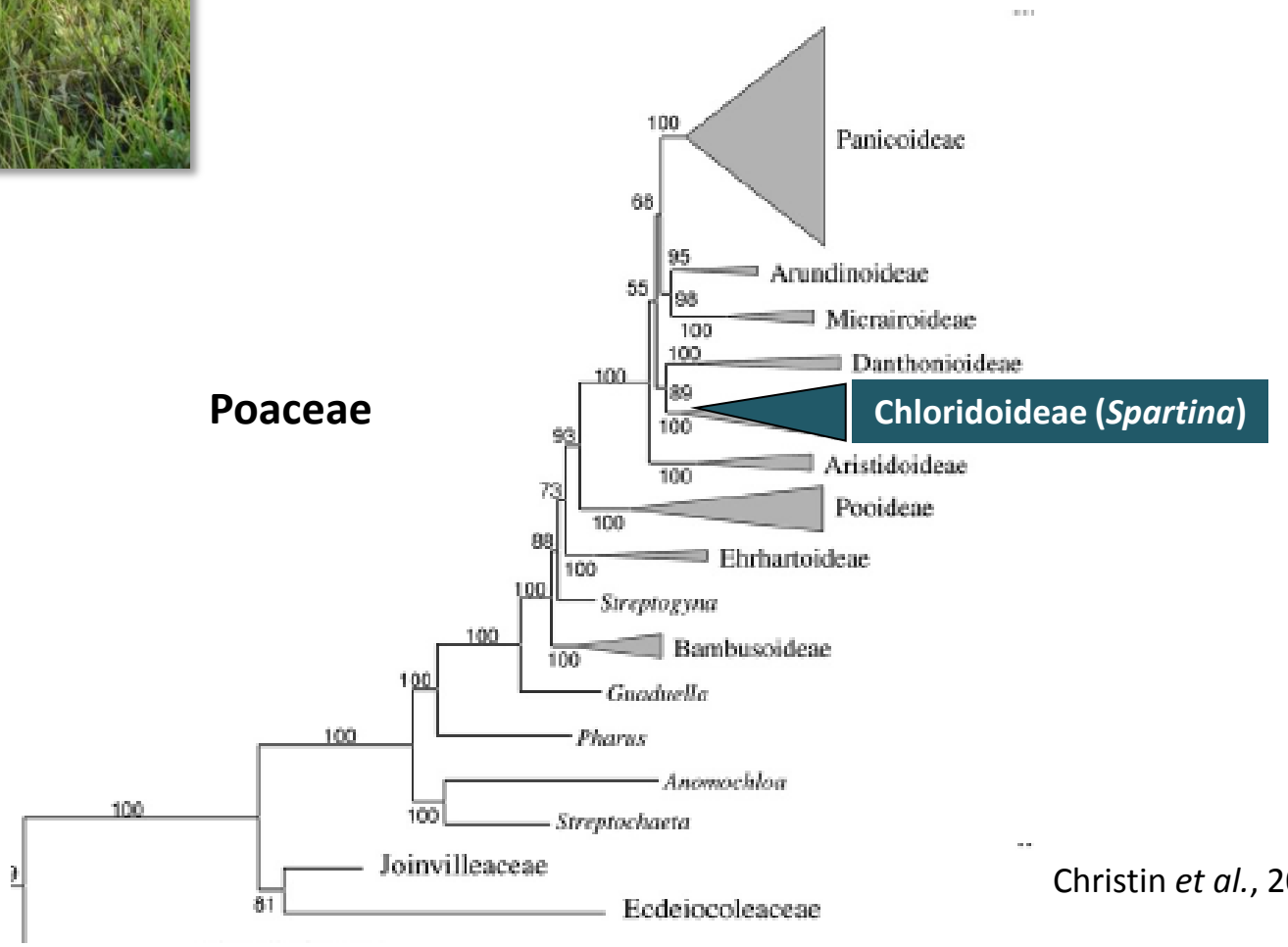
[1]CIRAD, UMR AGAP, Montpellier, France; [2]Université de Rennes 1, UMR CNRS 6553, Rennes, France; [3]INRA, CNRGV, Castanet Tolosan, France
[4]Genoscope, Evry, France

Réunion annuelle du Groupe Polyploïdie-Cytogénétique

Clermont-Ferrand, France, 22 mai 2014

Chloridoideae: a poorly investigated lineage in the Poaceae family

Poaceae

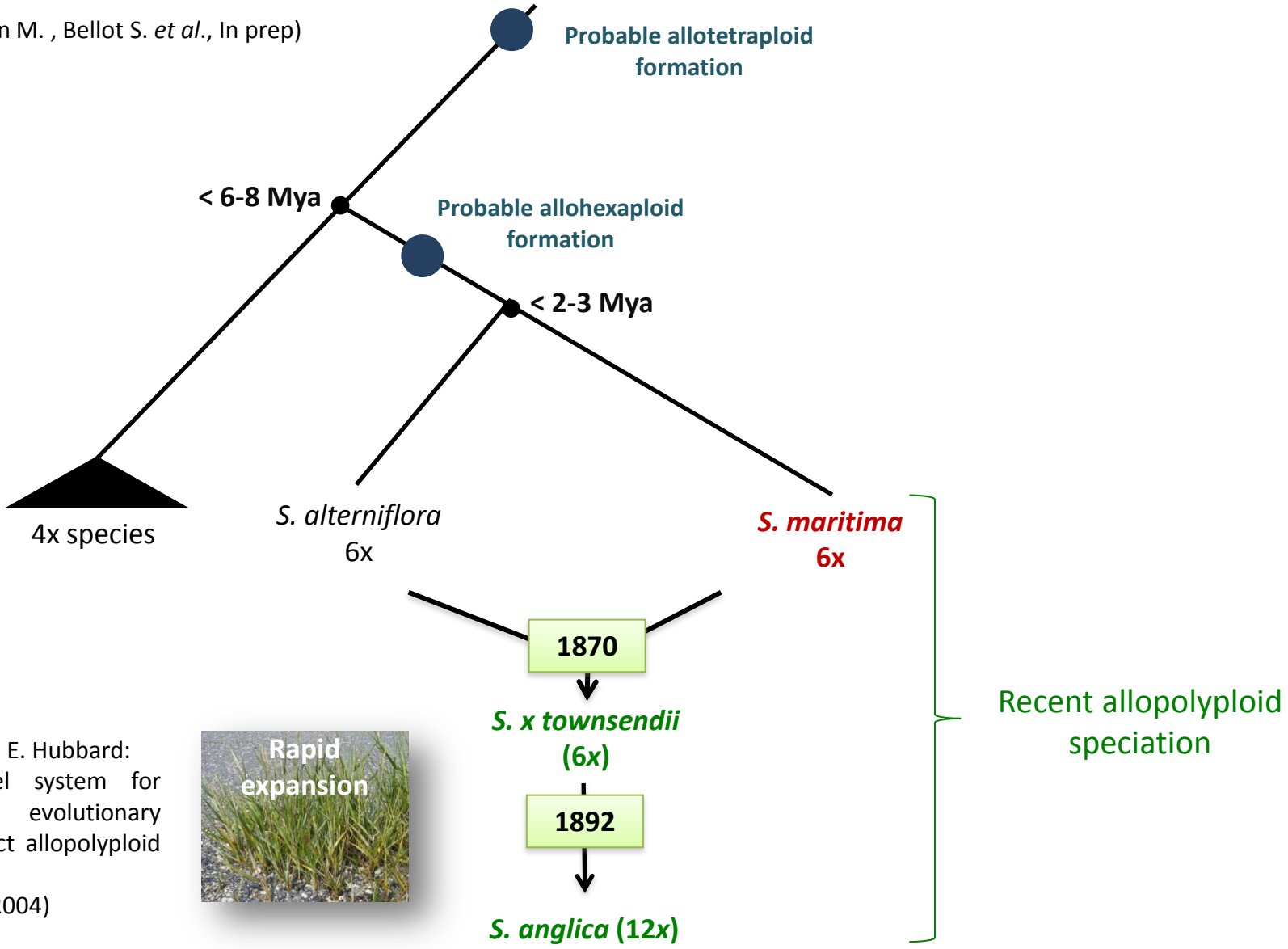


?

Christin *et al.*, 2009

Spartina: Recurrent hybridization and polyploidy

(Rousseau-Guetin M. , Bellot S. *et al.*, In prep)



Spartina anglica C. E. Hubbard: A natural model system for analysing early evolutionary changes that affect allopolyploid genomes.
(Ainouche *et al.*, 2004)

Spartina maritima: Evolutionary and ecological interests

2n = 6x = 60
2C = 3,7 pg, *c.a.* 3.7 Gb

Old-World native

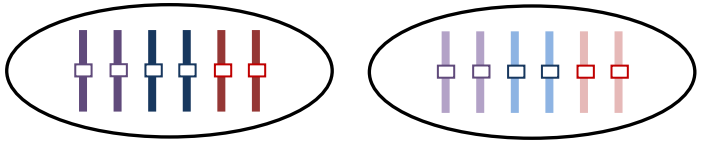
Distribution:
- Species in regression
- Local distribution
- Slow lateral extension

Biological and genetic diversity:
- Low seed-set
- Vegetative propagation
- Genetically depauperate populations
- Male parent of *S. x townsendii* et *S. anglica*

The *Spartina* challenge: no diploid reference genome

Strategy:

Analysis of homoeologous in the hexaploid parents



S. maritima *S. alterniflora*

- **Development of *S. maritima* genomic resources**

Reference transcriptome assembly
c.a. 17,000 annotated genes
Ferreira *et al.*, 2013

Genomic DNA from 454 Roche pyrosequencing
c.a. 100 Mb (A. Salmon)

BAC library construction
40,641 high-quality BAC-End Sequences
Ferreira *et al.*, 2013

First overview of the genomic composition in terms of gene content and repetitive sequences

- **Micro-analysis of homoeologous sequences, two complementary approaches :**

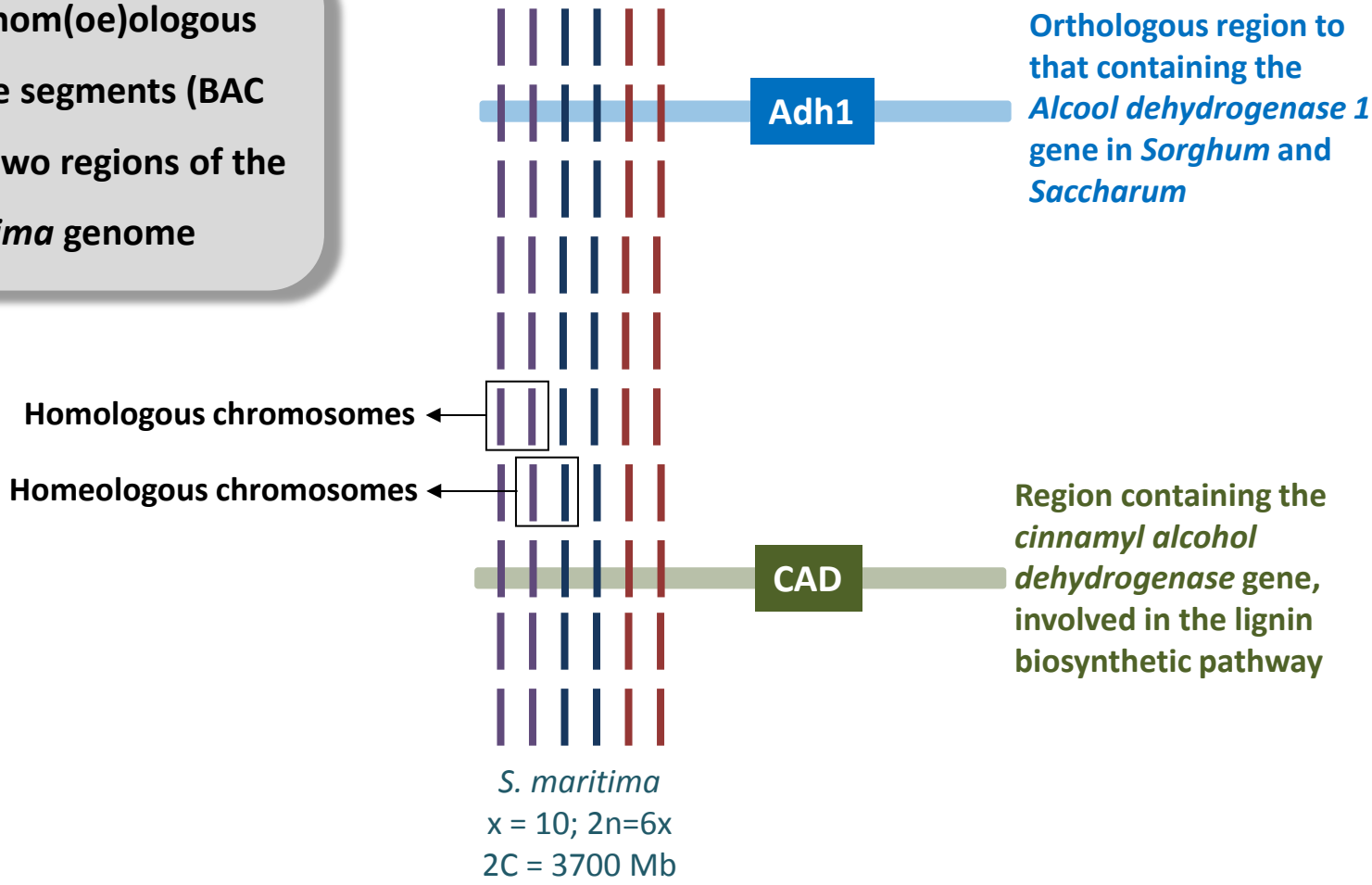
- Sequencing of a set of homoeologous BAC clones

- Intragenomic SNP analyses (NGS data): Julien Boutte, Malika Ainouche and Armel Salmon

Objectives

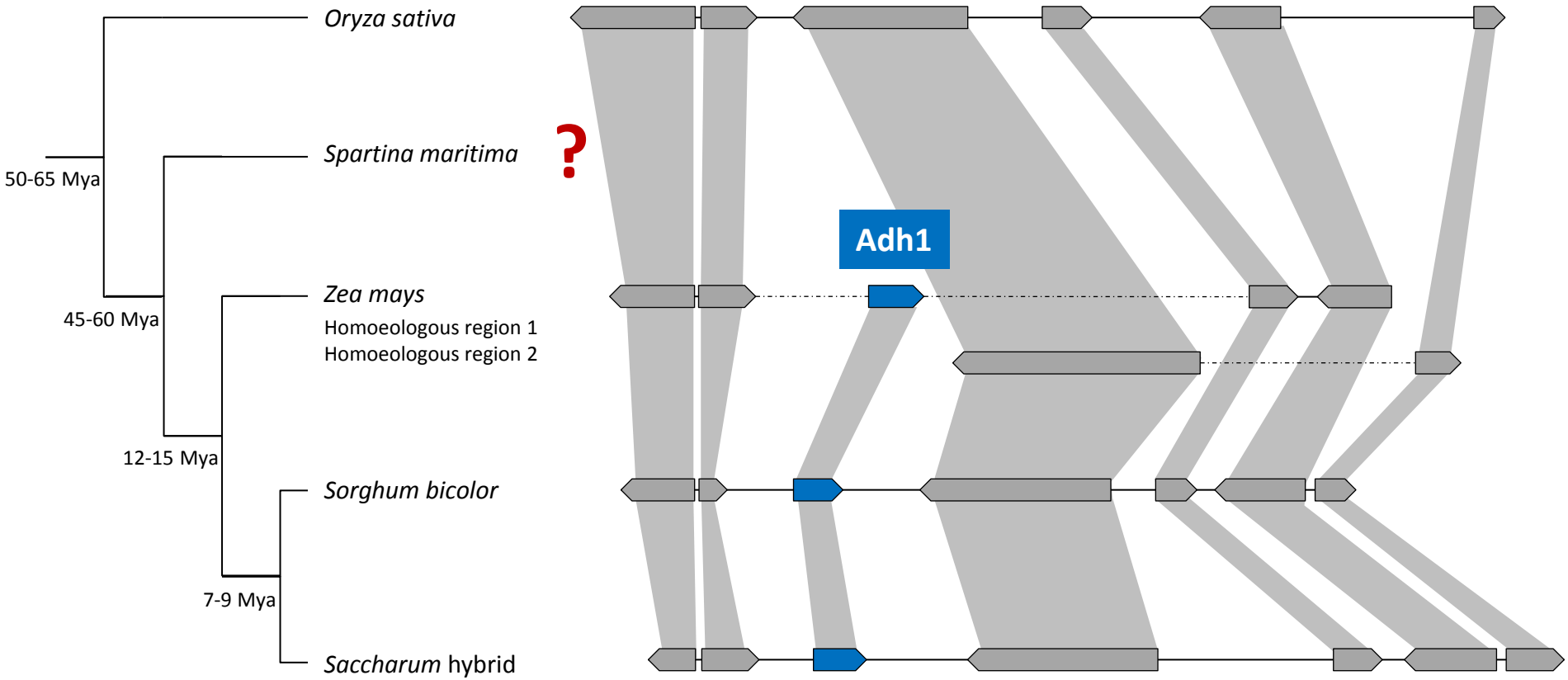
- To analyze homoeologous regions in the hexaploid genome
- To perform comparative analyses between *Spartina* (Chlorideae) and representatives from other grass subfamilies

Analysis of hom(oe)ologous chromosome segments (BAC clones) from two regions of the *S. maritima* genome



Adh1 region within Poaceae

Translocation containing *Adh1* gene previously detected in the Panicoids (*Sorghum*, *Saccharum* and *Zea*)



San Miguel et al., 1996 – Science; Tikhonov et al., 1999 – PNAS; Ilic et al., 2003 – PNAS; Jannoo et al., 2007 – Plant Journal

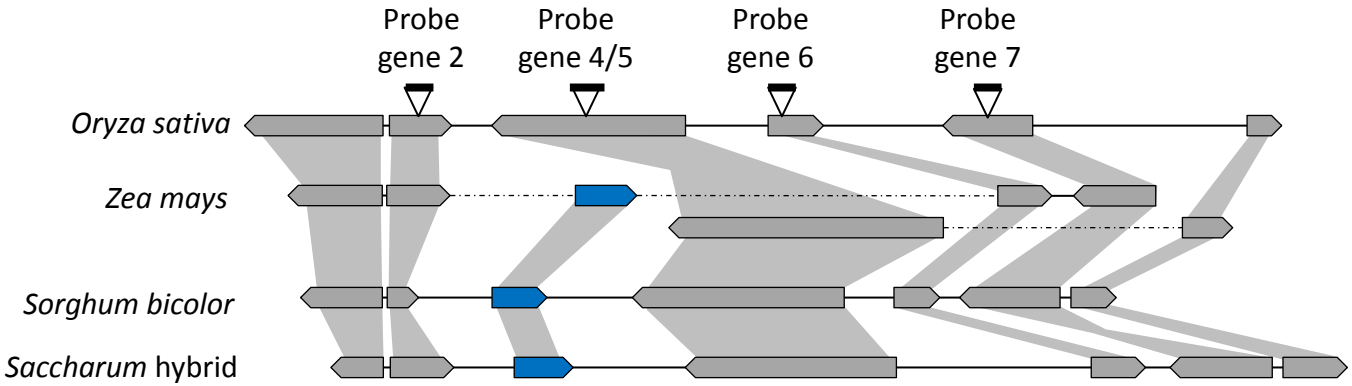
S. maritima BAC sequencing analyses

1. Identification of hom(oe)ologous BAC clones

S. maritima BAC library
Coll. H. Berges (CNRGV Toulouse)

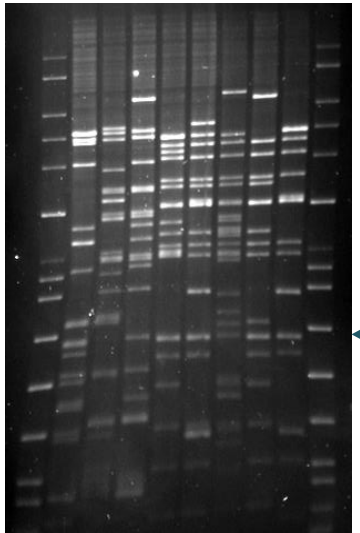
70 000 BAC clones (*c.a.* 110 kb) => 7.7 Gb
= 2x coverage of hexaploid genome

Probe design



BACs hybridizing to a minimum of two probes were identified and extracted

Hybridization of probes corresponding to the target region



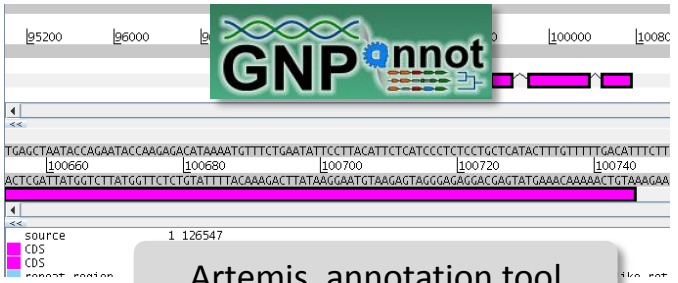
Fingerprinting to check BACs

S. maritima BAC sequencing analyses

2. **Sequencing and assembly of BAC clones** – Genoscope

3. **Automatic and manual annotation of BAC clones**

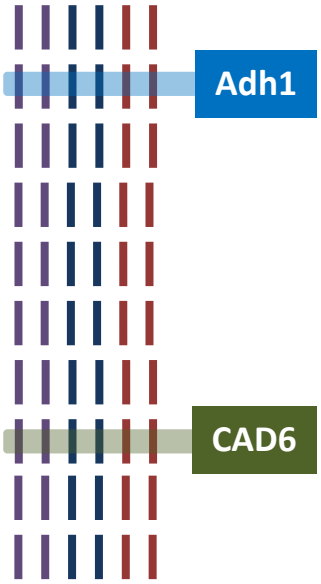
http://southgreen.cirad.fr/



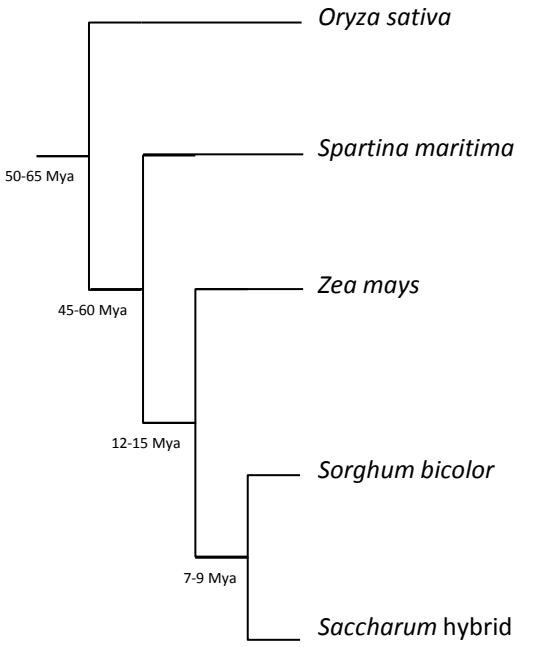
Artemis, annotation tool

4. **Intra- and inter-species comparisons:**

Between hom(oe)ologous *S. maritima* sequences



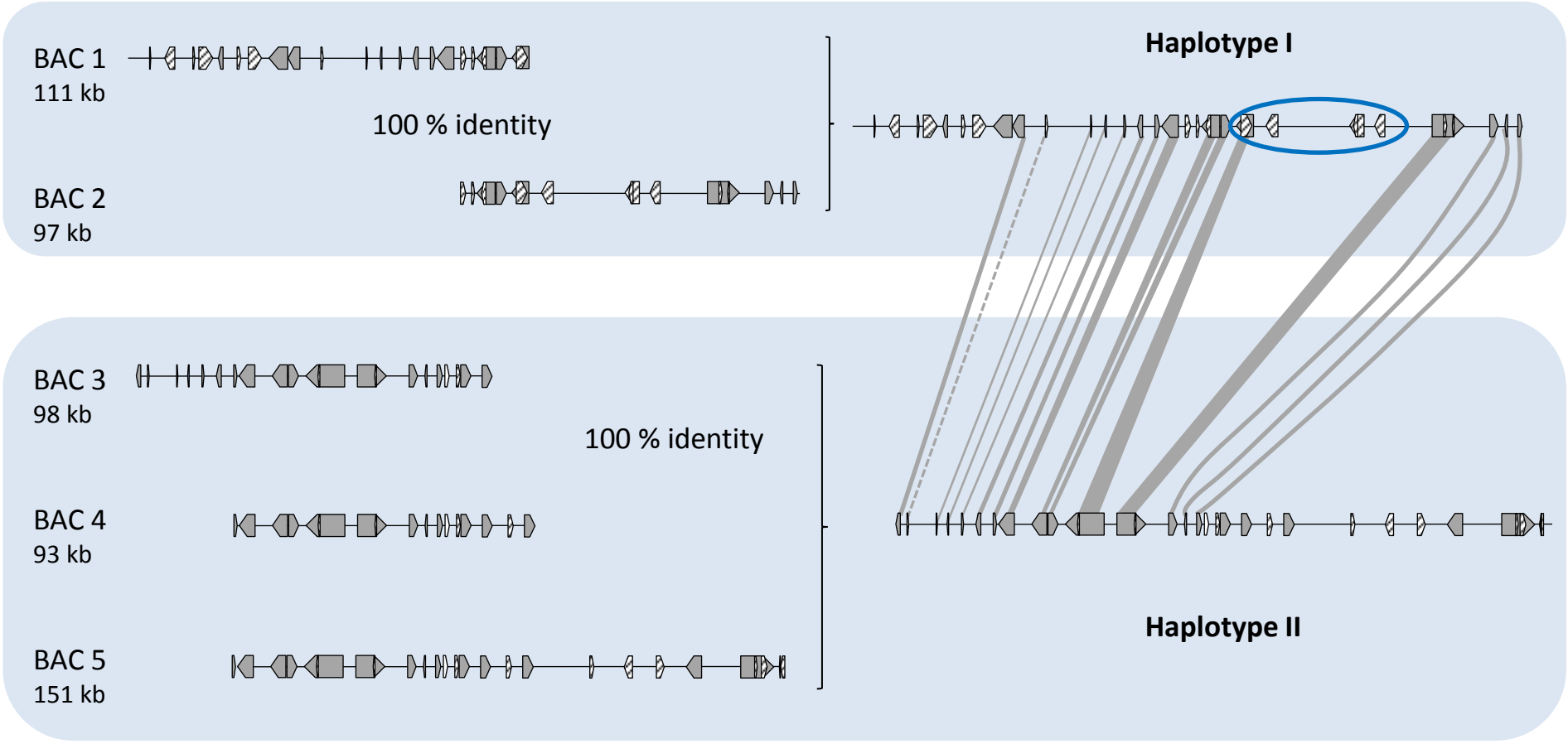
With orthologous Poaceae region



Adh1 region: Two divergent homoeologs identified

	Exons	Introns
% identity between homeologous genes	96 %	91 %

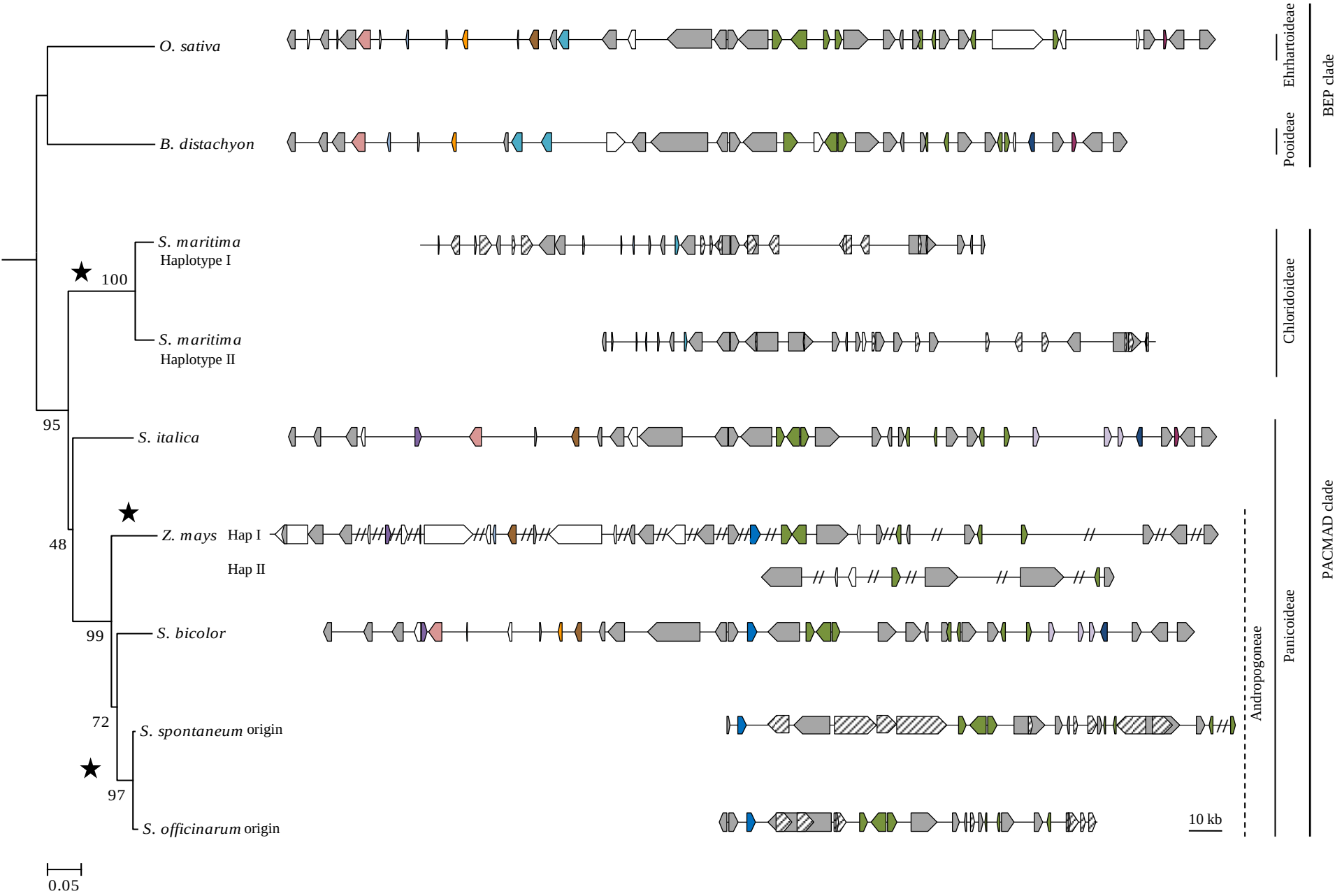
S. maritima homeologs



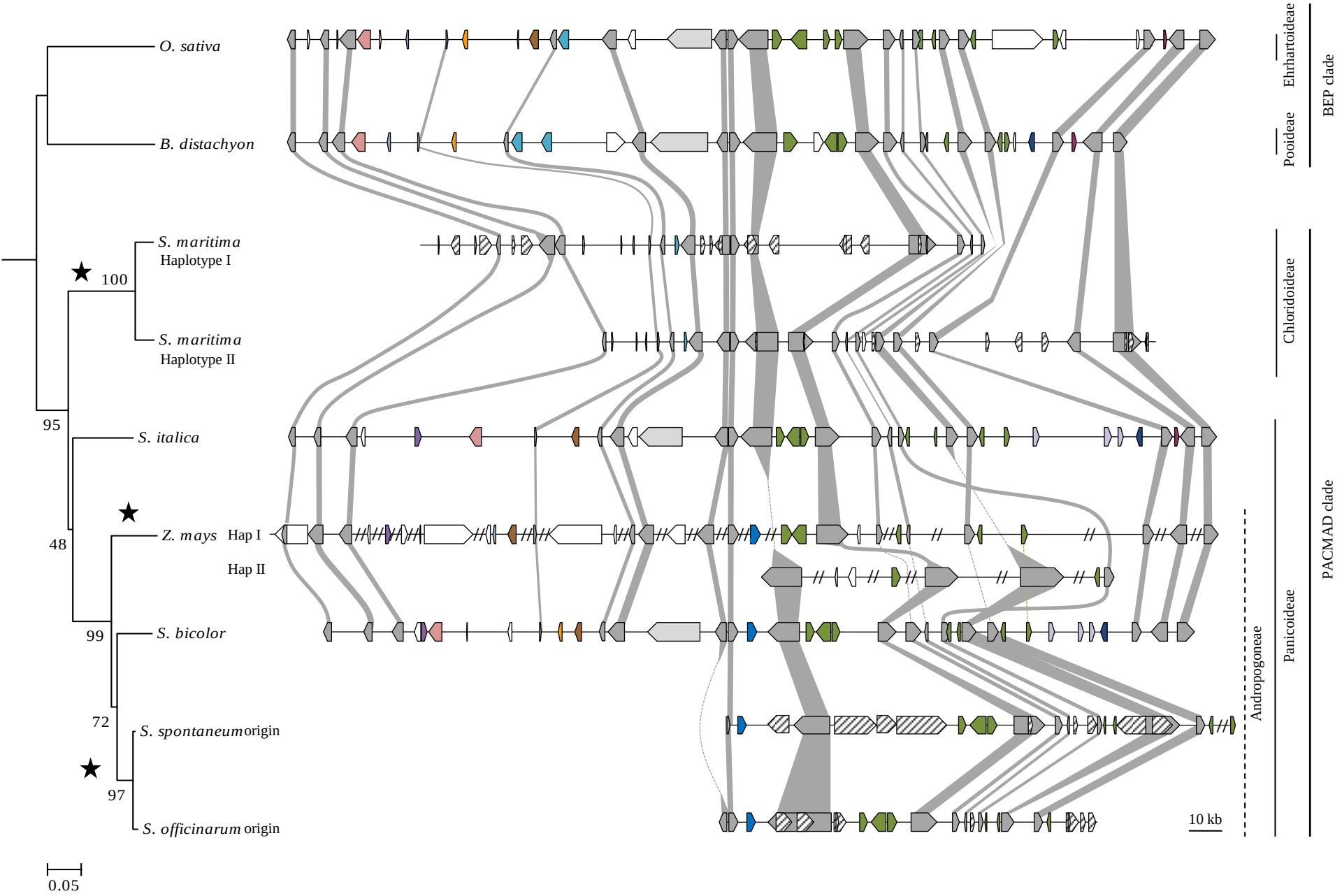
Genes
Collinear homeologous genes
Transposable elements (TEs)
Collinear homeologous TE

Differential TE insertion pattern between the 2 divergent homoeologous sequences (haplotypes I and II).

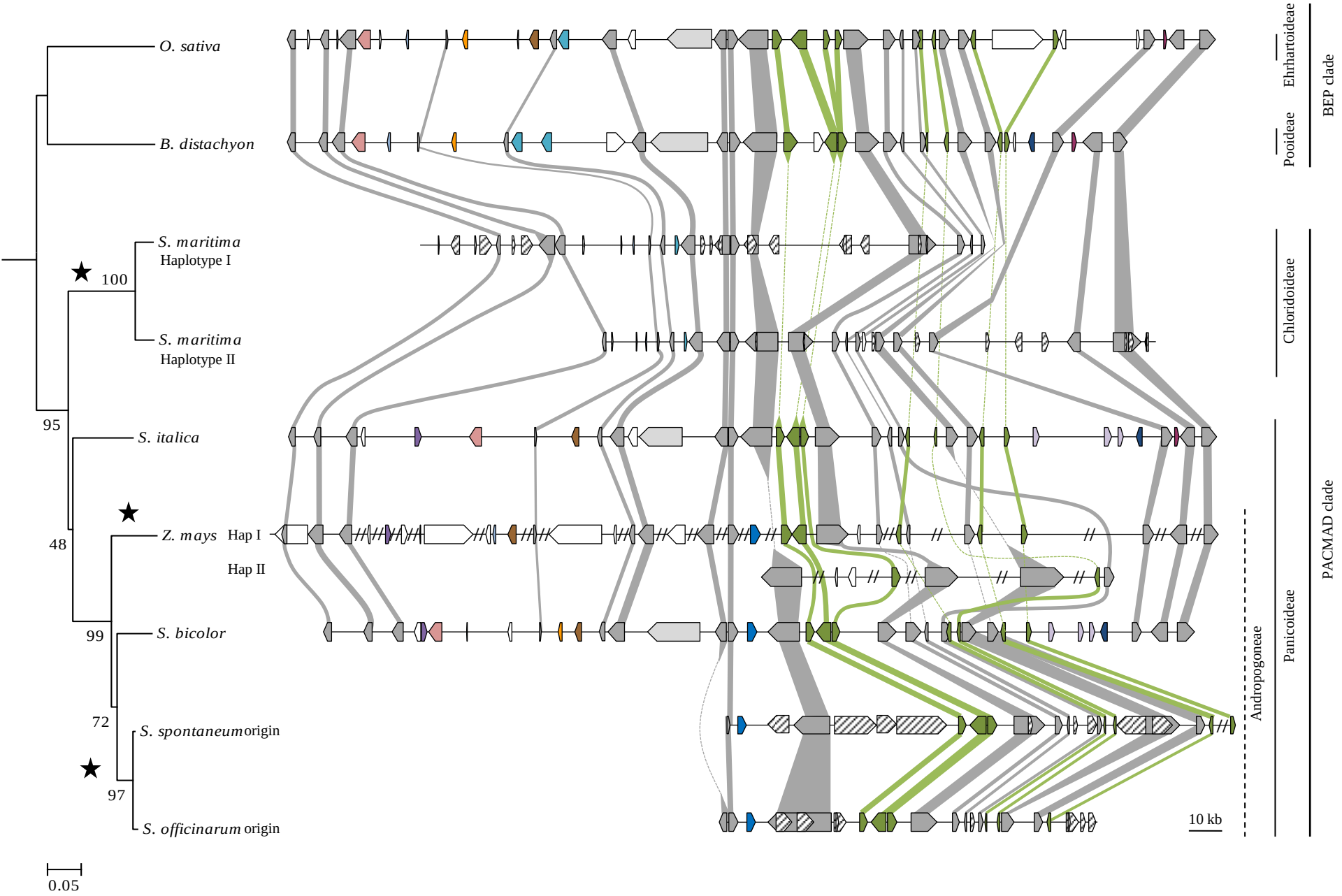
Adh1 region - Comparison with orthologous Poaceae chromosome segments



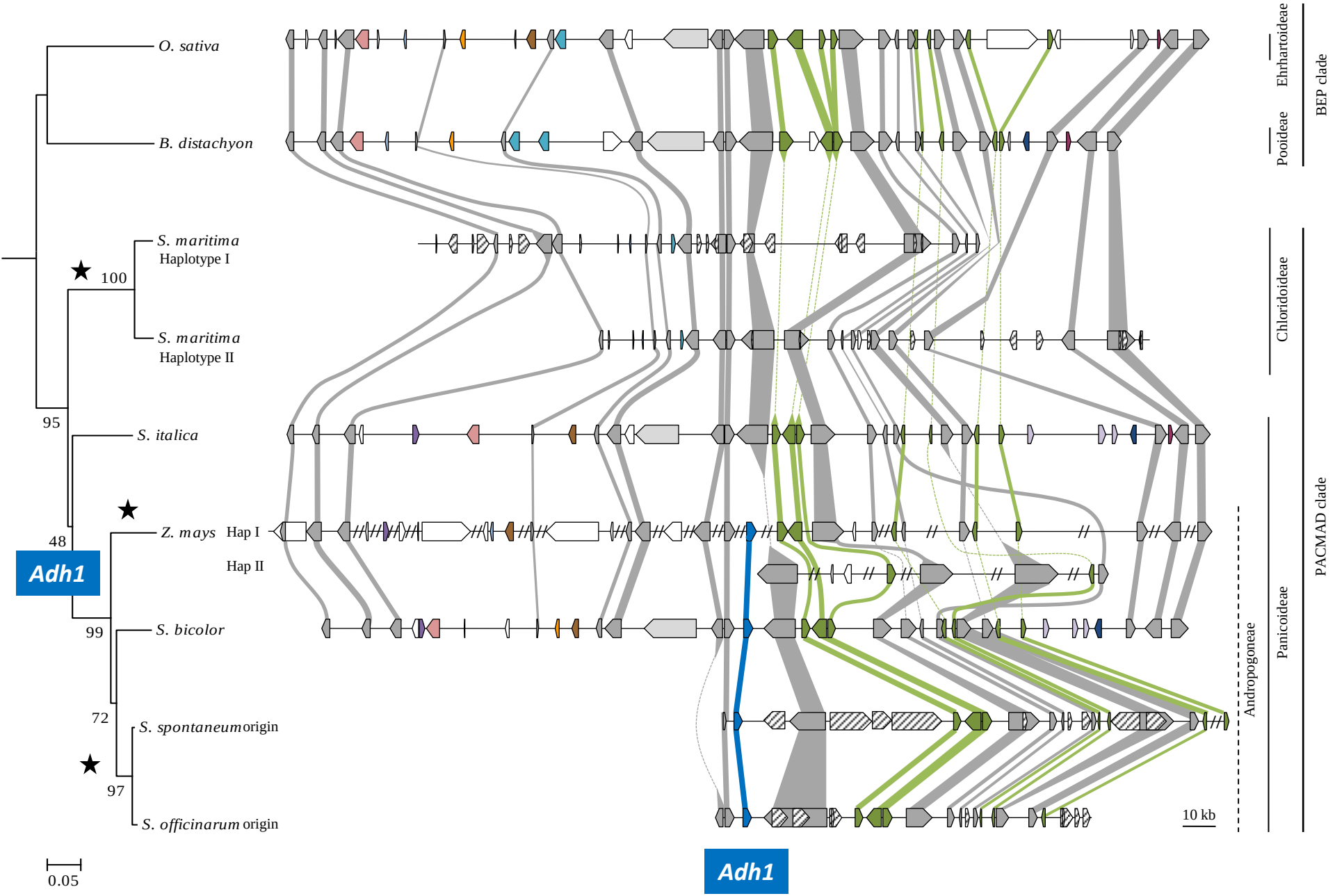
Adh1 region – 18 genes in all Poaceae chromosome segments



Adh1 region – 7 genes absent in *S. maritima* sequences



Adh1 region – no *Adh1* gene in *S. maritima* sequences



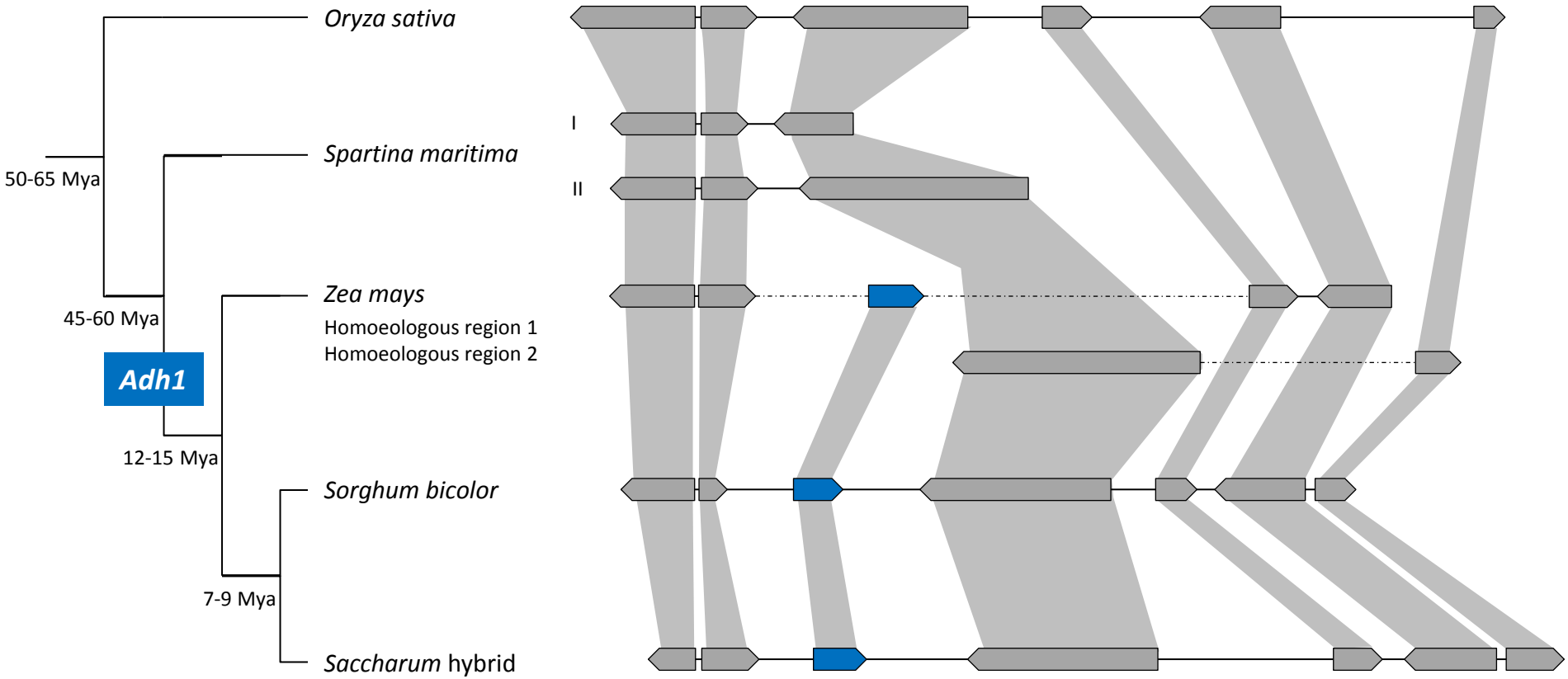
Adh1 region – Translocation after divergence Panicoicoideae/Chloridoideae

- Translocation containing *Adh1* gene previously detected in the Panicoids (*Sorghum*, *Saccharum* and *Zea*)

Jannoo et al., 2007; Ilic et al., 2003; Tikhonov et al., 1999, San Miguel et al., 1996

- Translocation not shared with *Spartina*

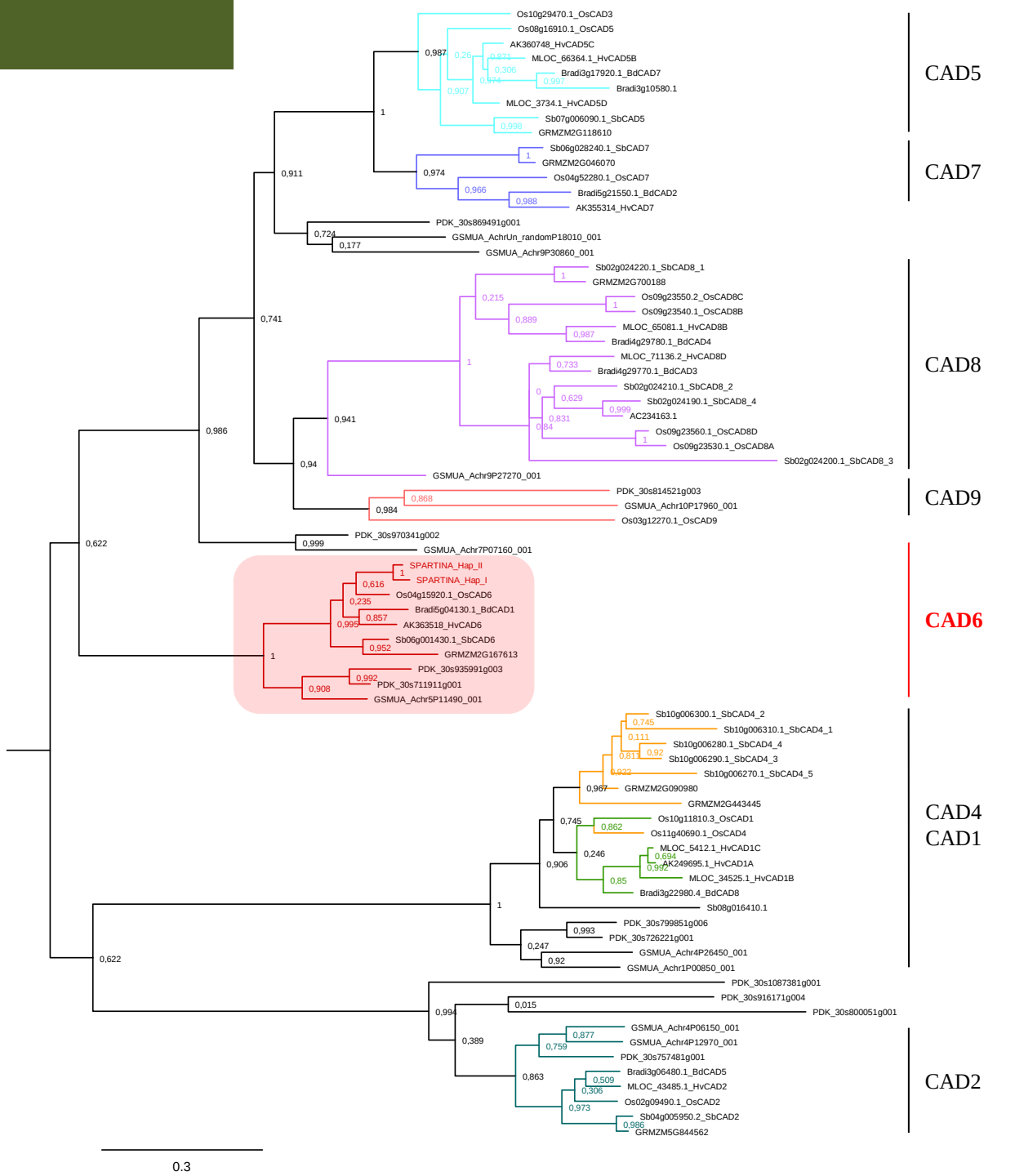
Translocation occurred after the divergence between Panicoideae and Chloridoideae 45-60 Mya



CAD region

Phylogenetic analysis of the CAD family in monocot plants

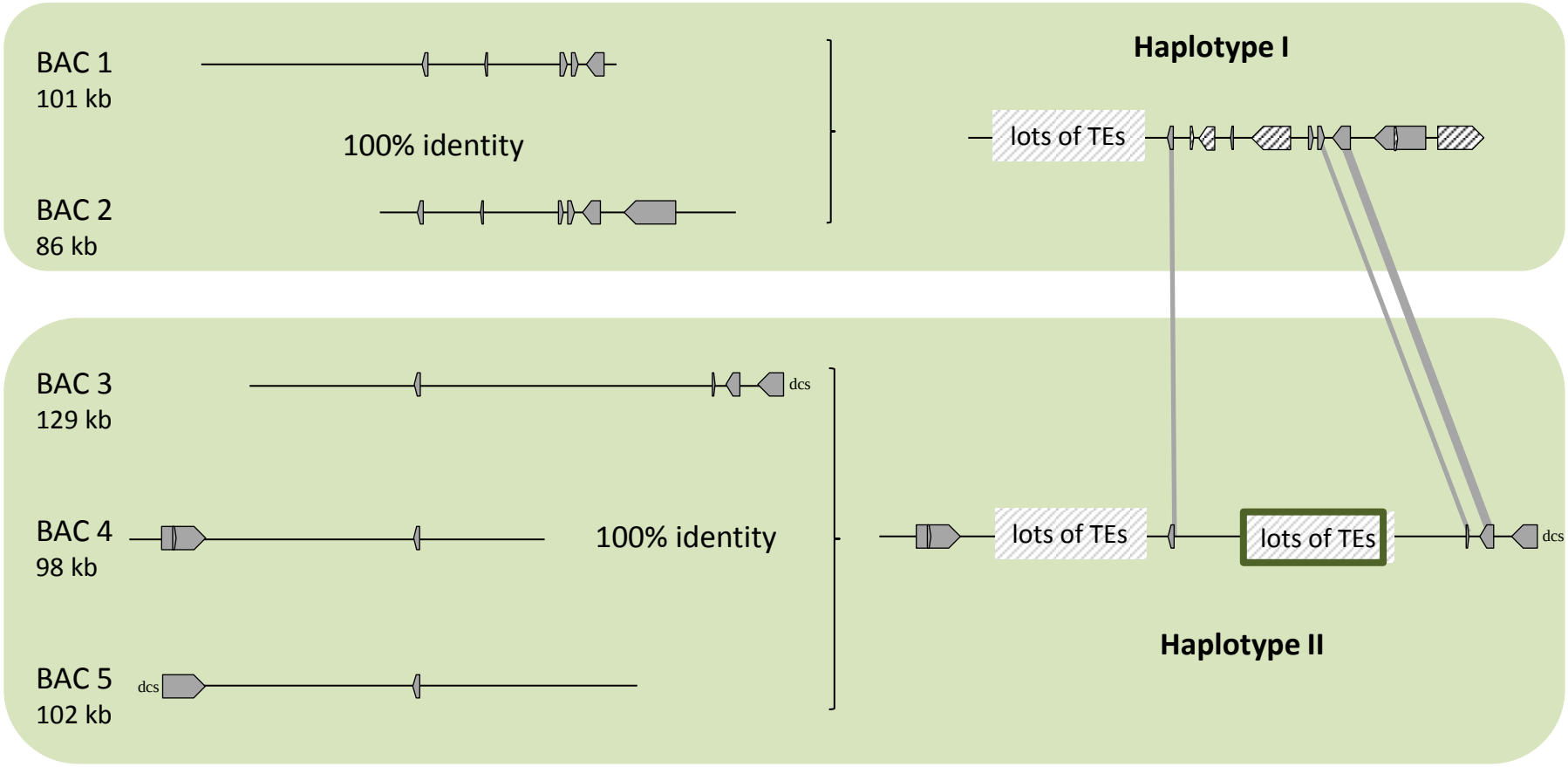
***S. maritima* region containing the *CAD6* gene**



CAD6 region : Two homoeologs identified

	Exons	Introns
% identity between homeologous genes	96 %	92 %

S. maritima homeologs



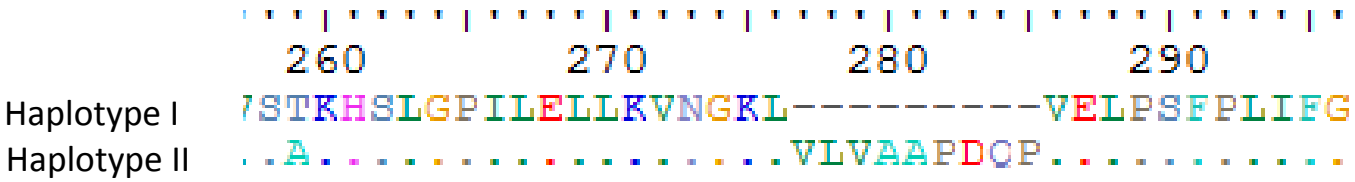
- Genes
- Collinear homeologous genes
- Transposable elements (TEs)

- **Structural rearrangements: 2 genes to only one of the two haplotypes**
- **Differential TE insertion pattern between the two haplotypes**

CAD6 region: Two homoeologs identified

Important divergence between the two CAD6 homoelogs:

Insertion/deletion of 9 amino-acids



What about evolution of the *CAD6* gene in other *Spartina* species?

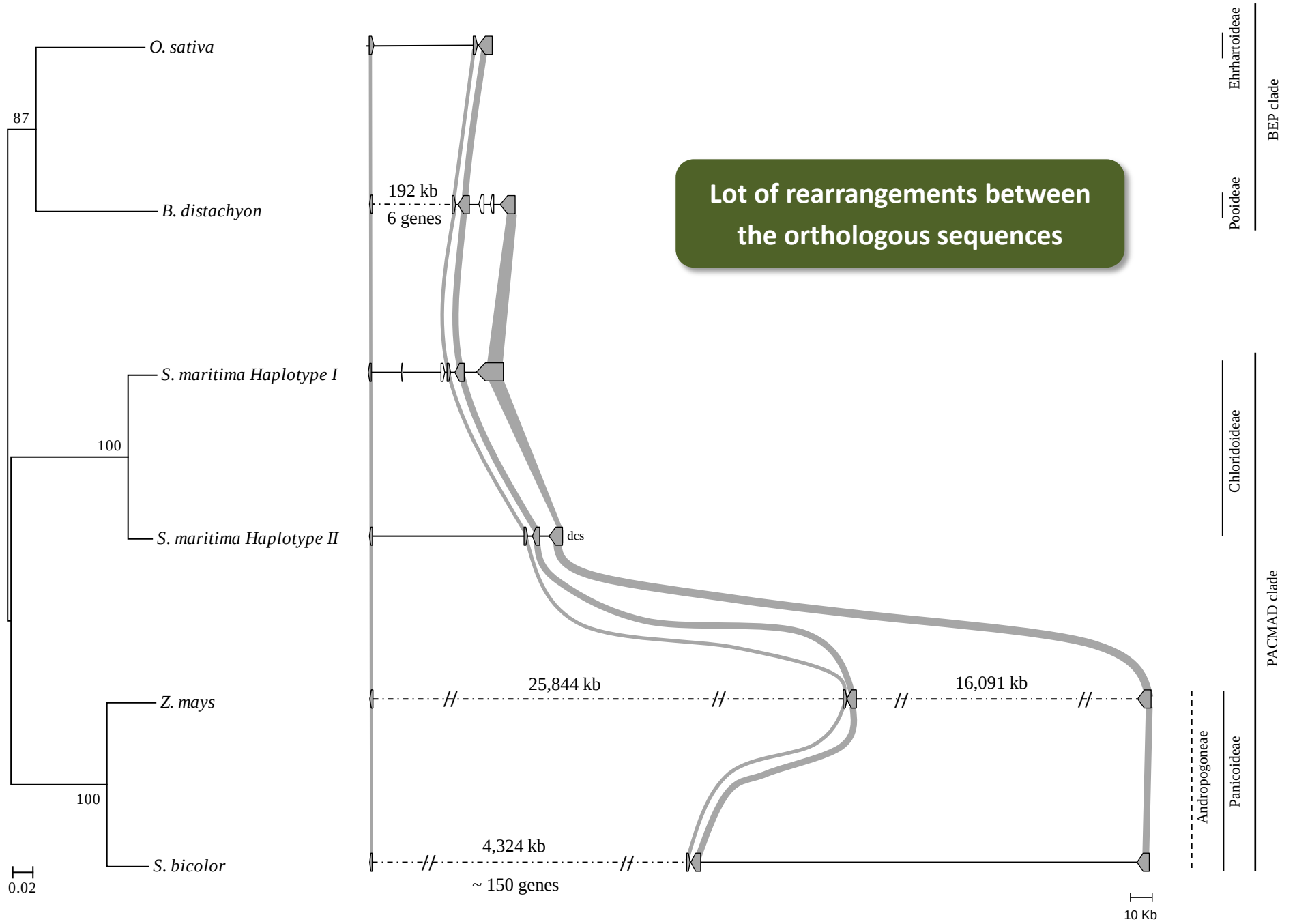
Next steps:

Primer design

PCR amplification on gDNA from other *Spartina* species

Phylogeny reconstruction based on CAD6 sequences

CAD6 region - Comparison with orthologous Poaceae chromosome segments



Evolution of the hexaploid *S. maritima*

Two divergent homoeologs detected in both regions

	% identity exons	% identity introns	Ka	Ks	Ka/Ks
Adh1 region	96	91	0.016	0.030	0.54
CAD6 region	96	92	0.023	0.058	0.47

But 3 homoeologs expected in hexaploids:

- Gene loss – fractionnation?
- Technical bias?
- Incomplete BAC sampling?

NGS data

***In silico* haplotype detection from transcriptomic NGS reads (454 + Illumina)**

13 genes from the Adh1 region screened

5 genes were retrieved (expressed in cDNA)

42/46 homoeosnps validated

7/10 were 100% validated haplotypes

from 3 to 18 haplotypes / alignement

Gene #13

Haplotype I NGS

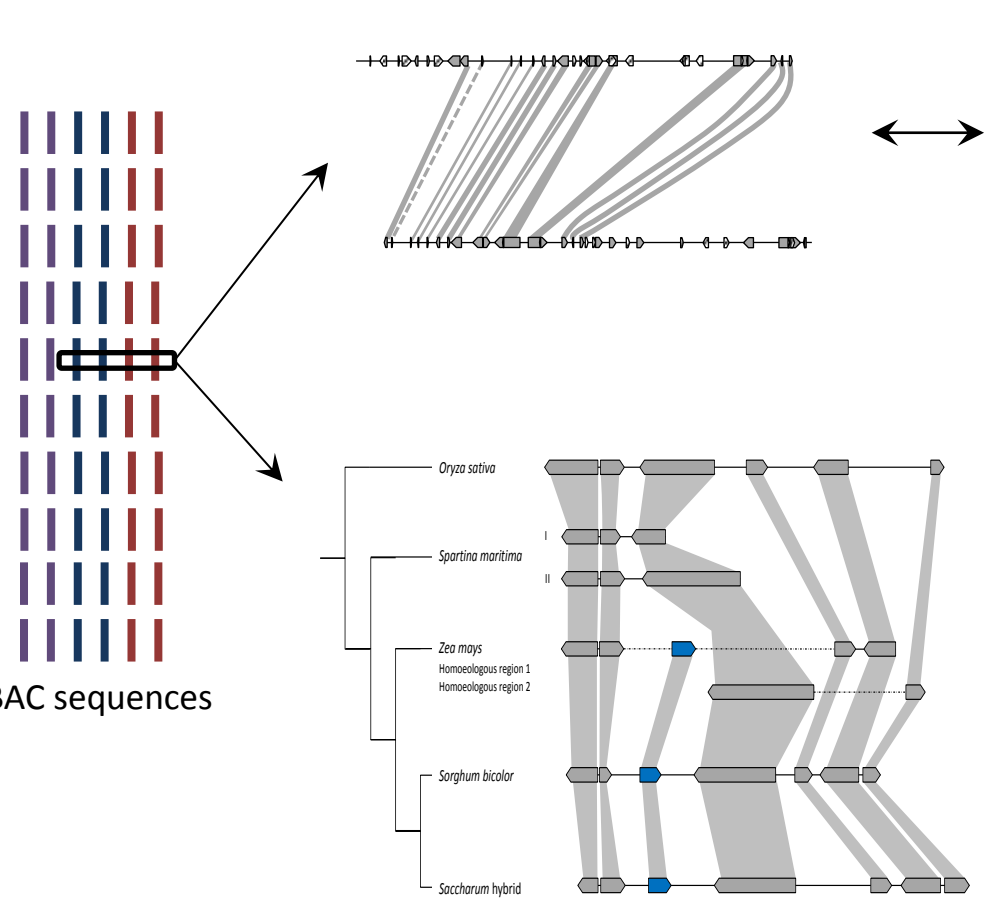
Haplotype II NGS

Haplotype III NGS

Boutte *et al*. in prep

Conclusion

First homoeologous genomic regions investigated in *Spartina*



BAC sequences

Large-scale NGS data

Insights into evolution of complex genomes in *Spartina* genus

MOB team
Mathieu Rousseau-Gueutin
Julie Ferreira de Carvalho
Houda Chelaifa
Julien Boutte
Armel Salmon
Malika Ainouche

http://ecobio.univ-rennes1.fr/

SEG team
Carine Charron
Angélique D'Hont

Bioinformatics team
Gaëtan Droc

http://umr-agap.cirad.fr

Arnaud Bellec
Joëlle Fourment
Hélène Bergès

http://cnrgv.toulouse.inra.fr/fr

Arnaud Couloux
Julie Poulain
Sophie Mangenot
Patrick Wincker

http://www.genoscope.cns.fr/spip/